

## 以無人機進行生態型市民農園 六種雜草影像分析及土壤情況分析

### Image Identification and Soil Condition Analysis of Six Weeds in Ecological Citizen Farms Using Drones

逢甲大學自動控制工程學系

蔡翔勛、陳彥綸\*、蘇子傑、林宸生

#### 中文摘要

因應現代工業 4.0 的發展，自動化運用於傳統農園對比過往來得進步許多，高科技也加入了農業進化之歷程，形成「智慧農園」。

本研究主旨使用 MATLAB 程式語言來實現深度學習。深度學習是以人工神經網路為架構的一種演算法，其中以 CNN 卷積神經網路架構最具代表性，其架構主要可分為卷積層、池化層及全連接層，及深度學習是藉由資料自我學習，並採用了深層類神經網路結構來訓練或分析資料，用來學習如何進行影像辨識及分類雜草。

收集雜草資料時，因雜草範圍過大，導致拍攝過程不便，因此採用無人機空拍之方式以照到更大範圍的雜草，再將無人機所拍攝到的照片，回傳至電腦進行影像分析顯示為何種雜草，此外，我們想進一步瞭解雜草在土壤裡的貧瘠及濕潤程度，結合了土壤溼度感測器檢測出實驗所需數據，並將即時結果顯示於人機介面中。

**關鍵字：**深度學習、影像辨識、無人機、感測器、人機介面

#### 一、研究動機與目的

生活中對於雜草的印象，莫過於「似無所見，又非視而不見」。雜草與我們的生活有著深厚的關係，到處都能看見它的蹤影。有的雜草會為我們帶來好處，有的卻會使我們丟了性命。基於此概念，我們突發其想，設計出能夠辨識身邊帶來正負面效益雜草之系統，進而瞭解其功用，更推廣至農業領域中成為不可或缺之實用工具。

本專題研究利用搭載攝像鏡頭之無人機結合影像辨識方式，針對雜草種類進行影像分析，無人機於農園裡拍攝雜草，將資料傳回電腦透過 MATLAB，經由深度學習之模型進行影像分析，迅速瞭解雜草種類；並於農園土壤中插入土壤溼度感測器測得溼度數值，傳入人機介面顯示於其中

隨著農業逐漸轉型，過程中自原先農業 1.0 之徒手種植，漸進式地加入了犂牛進行耕種，接著工具逐漸升級以車替代；近年來，高科技也加入了農業進化之歷程，形成「智慧農園」。智慧農園源於 104 年起行政院農業委員會推動之「農業 4.0」發展方案，期待以感測器、智能裝置、物聯網、資料庫，

建構智農產銷與數位服務體系；更藉由生物科技、資料通訊科技與自動化機械的規模生產，精準提升產值。

本研究以遠端程式操控無人機於各地飛行拍攝，畫面中可觀察到包羅萬象之植物及雜草，並將拍攝內容回傳至 MATLAB 進行雜草種類多廣的辨識；再於人機介面上分析土壤感測器回傳之資料，產出此地適合哪種雜草生存的結果。使農民於選擇種植地上的效率大大提升，也不用頂著大太陽一一觀察雜草種類，更能有效預防被昆蟲咬傷、分泌物灼傷等問題。

總結來說，結合無人機辨識，進行生態型市民農園雜草影像分析與土壤情況分析，能夠節省人力資源並達到風險預防之目的，精確度更高於長年人工觀察之結果。切合智慧農園之推動方向，我們將帶領農業邁向與數位科技時代合作之雙贏局面！

## 二、研究方法

### 2.1 機器學習與深度學習比較

近幾年人工智慧(AI)[1]迅速的發展，源於 AlphaGo 在圍棋上不斷擊敗人類，因而 AI 聲名大噪，其實在 1956 年人工智慧這個名詞便已出現，其簡單定義為讓電腦模擬人類做的事情；機器學習(Machine Learning)[2]是 AI 的一部分，也是實現 AI 的其中一種方法；深度學習(Deep Learning)[3]則是涵蓋在機器學習裡，是機器學習的其中一個方法。圖 2.1 為人工智慧、機器學習、深度學習的演進。

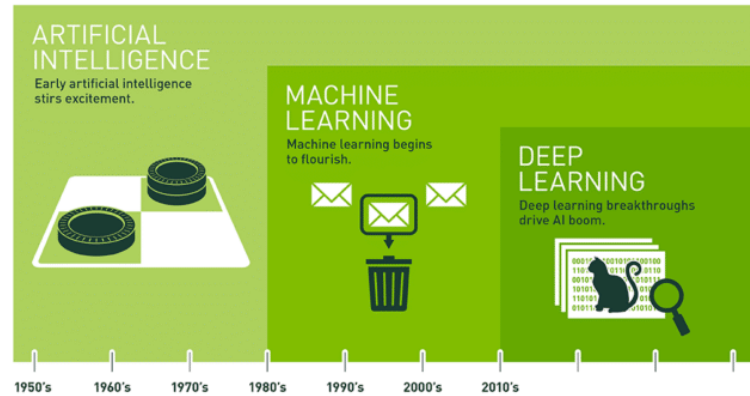


圖 2.1 人工智慧、機器學習與深度學習的演進

### 2.2 CNN 介紹

卷積神經網路 CNN [4]，是一種前饋神經網路(如圖 2.2 為 CNN 卷積神經網路)，它的人工神經元可以回應一部分覆蓋範圍內的周圍單元，對於大型圖像處理有出色的表現。卷積神經網路是由一個或多個卷積層和全連接層組成，同時也包括關聯權重和池化層(pooling layer)。這一結構使得卷積神經網路能夠利用輸入資料的二維結構，與其他深度學習結構相比，卷積神經網路在圖像和語音辨識方面能夠給出更好的結果。這一模型也可以使用反向傳播演算法進行訓練。相比較其他深度、前饋神經網路，卷積神經網路需要考量的參數更少，使其成為一種具有吸引力的深度學習結構。

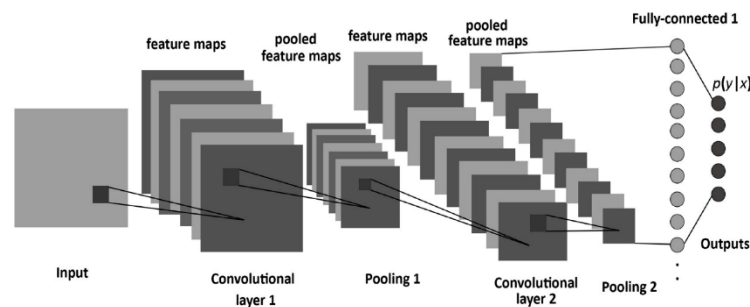


圖 2.2 CNN 卷積神經網路

## 2.3 AlexNet 介紹

AlexNet[5]在深度學習領域中具有相當重要的地位，由 Alex Krizhevsky 於 2012 年時所提出，並且於同年的 ImageNet LSVRC 競賽中奪得了冠軍。而 AlexNet 的架構主要是包含了 5 層的卷積層、3 層的全連接層(圖 2.3 為 AlexNet 架構)，並且在第六層及第七層中加入了 Dropout，使得模型不會過度依賴於某些特徵，進而增強了模型的泛化能力，表 2.1 為 AlexNet 主要特點。

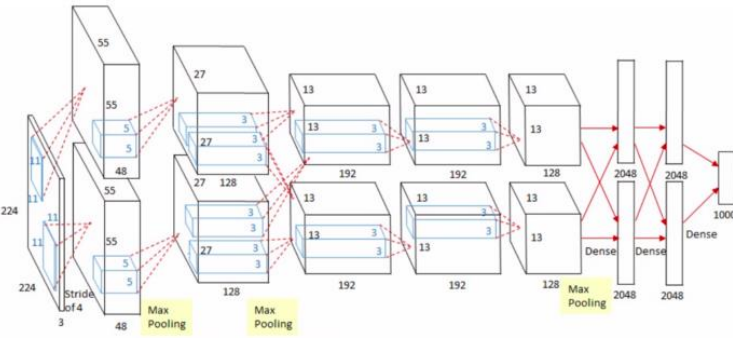


圖 2.3 ALEXNET 架構

表 2.1 AlexNet 主要特點

- 使用數據增強方式來抑制模型產生過擬合的問題
- 使用 Relu 作為非線性函數，來避免因為神經網路層數過深或梯度過小，所造成的梯度消失問題
- 使用多個 GPU 來訓練模型
- 使用層疊的卷積層來提取特徵
- 使用 Dropout，來防止模型產生過擬合
- 將輸入層數變得更大，可以輸入 224x224 尺寸大小的彩色圖片

## 2.4 ResNet50 介紹

ResNet(Deep Residual Networks) 稱為深度殘差網路(圖 2.4 為 ResNet 架構)，是一種利用殘差學習的神經網路 [6]，而層數高達 152 層，藉由加入大量 Shortcut 來提升模型的精確度，並且也解決了模型發生梯度爆炸、梯度消失等問題，而在模型上 ResNet-18、ResNet-34 都採用了 building block，ResNet-50、ResNet-101 則採用 bottleneck build block，目的是為了降低模型的複雜度。

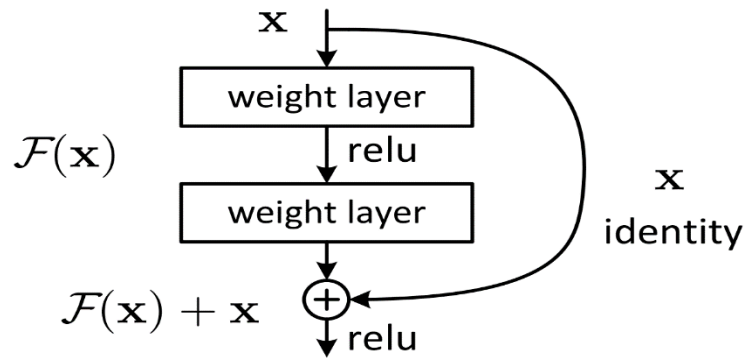


圖 2.4 ResNet 架構

## 2.5 VGG 介紹

在 2014 年的 ImageNet 競賽中，VGGNet 想要探討卷積神經網路的深度與其性能之間的關係，比起 AlexNet 採用更深層的網路，其特點是重複採用同一組基礎模組使整體架構看起來更整潔，並使用小卷積核取代 AlexNet 裡的中大型卷積核(圖 2.5 為採用小卷積核取代中大型卷積核示意圖)，其架構由 N 個 VGG Block 與 3 個全連接層所組成，VGG Block 的構造就是由不同數量的 3x3 卷積層，以及 2x2 的最大池化層所組成，最終獲得 ImageNet 競賽的亞軍。

VGGNet有許多不同的結構，例如VGG11、VGG13、VGG16[7]、VGG19[8]，其差異在於網路的層數(卷積層與全連接層的數量)。常見的結構為VGG16(圖2.6為VGG16架構)，其架構使用了5個卷積層與3個全連接層，其中前兩個卷積層內含2個基礎模組、後三個卷積層內含3個基礎模組，總共為 $2 \times 2 + 3 \times 3 + 3 = 16$ 層網路層數。表2.2為VGG模型優缺點。

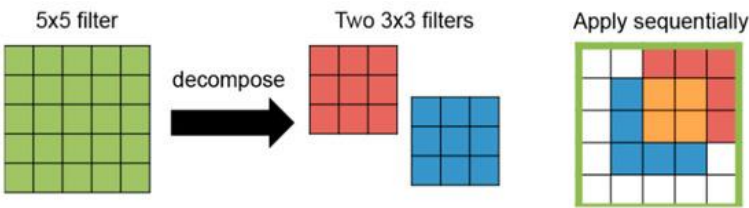


圖 2.5 小卷積核取代中大型卷積核示意圖

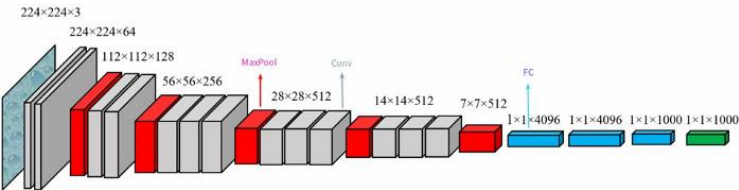
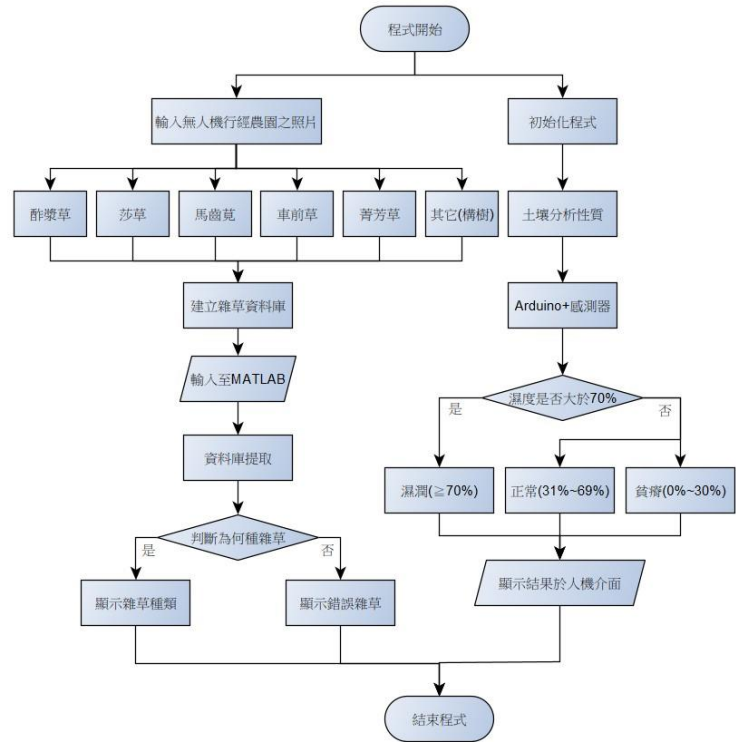


圖 2.6 VGG16 架構

表 2.2 VGG 模型優缺點

優點	<ul style="list-style-type: none"> <li>●結構簡潔，整個構造使用同樣的大小的卷積層和最大池化層</li> <li>●使用多個較小的卷積核的組合比一個大的卷積核來的好</li> <li>●透過不斷加深網路結構可以提升性能</li> </ul>
缺點	<ul style="list-style-type: none"> <li>●耗費更多計算資源，並且使用了更多的參數，導致更多的內存占用</li> </ul>

### 三、系統流程



### 四、實驗結果與分析

#### 4.1 影像前處理

##### 4.1.1 照片九宮格

在拍攝照片時一定會再拍到周圍的東西，也因此周圍的東西會影響辨率，所以我們上網搜尋切割照片之Python程式，進而加以修改成照片原圖切割九宮格取中間照片之程式以達到不降低辨識率之目的，圖4.1為切割九宮格流程。

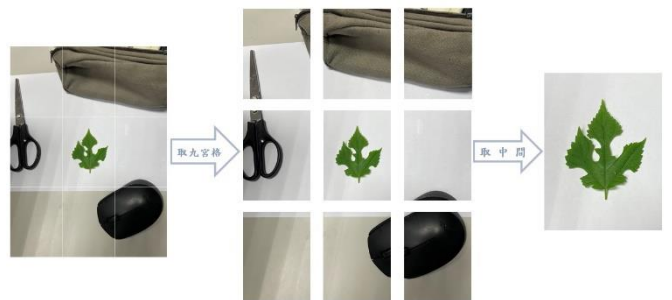


圖 4.1 切割九宮格流程

#### 4.1.2 圖像增強-旋轉及展延

圖片在輸入後只有當初拍攝的單一角度，為了讓電腦在深度學習時能夠學習照片的更多角度以及展延後照片的多樣化，所以利用 MATLAB 的圖像增強功能，讓辨識時能有更高的辨識率。圖 4.2 為旋轉及展延程式數值。

##### ●RandRotation-旋轉範圍(預設為[0 0])

應用於輸入圖像的旋轉範圍（以度為單位），指定為以下值之一。

- 兩元素數值向量。第二個元素必須大於或等於第一個元素。旋轉角度是從指定間隔內的連續均勻分佈中隨機選取的。

- 函數句柄(function handle)。該函數不得接受任何輸入參數，並以數值標量形式返回旋轉角度。使用函數句柄從不相交的區間或使用非均勻概率分佈選取旋轉角度。

- 預設的情況下，不旋轉增強圖像。

##### ●RandXTranslation-水平平移範圍(預設為[0 0])

應用於輸入圖像的水平平移範圍，指定為以下之一。平移距離以像素為單位。

- 兩元素數值向量。第二個元素必須大於或等於第一個元素。水平平移距離是從指定間隔內的連續均勻分佈中隨機選取的。

- 函數句柄(function handle)。該函數不得接受任何輸入參數，並將水平平移距離作為數值標量返回。使用函數句柄從不相交的區間或使用非均勻概率分佈選取水平平移距離。

- 預設情況下，增強圖像不會在水平方向上平移。

##### ●RandYTranslation-垂直平移範圍(預設為[0 0])

應用於輸入圖像的垂直平移範圍，指定為以下之一。平移距離以像素為單位。

- 兩元素數值向量。第二個元素必須大於或等於第一個元素。垂直平移距離是從指定間隔內的連續均勻分佈中隨機選取的。

- 函數句柄(function handle)。該函數不得接受任何輸入參數，並將垂直平移距離作為數值標量返回。使用函數句柄從不相交的區間或使用非均勻概率分佈選取垂直平移距離。

- 預設情況下，增強圖像不會在垂直方向上平移。

```
img_size=30; %input image size
imageAugmenter = imageDataAugmenter( ...
    'RandRotation',[-5,5], ... %旋轉角度
    'RandXTranslation',[-2 2], ...%X軸的展延
    'RandYTranslation',[-3 3]); %y軸的展延
```

圖 4.2 旋轉及展延程式數值

### 4.2 訓練過程波形與過擬合現象

訓練進行中會有視窗可以監控訓練時的當下狀況，我們得以去判斷訓練是否發生過擬合現象或是預測辨識率不理想等等之情況，圖 4.3 為訓練過程。

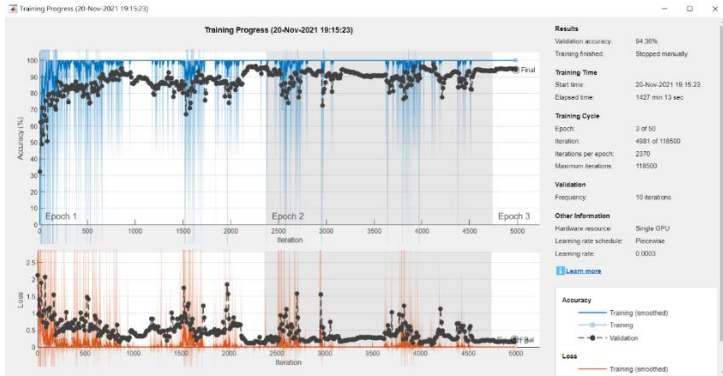


圖 4.3 訓練過程

模型對訓練集數據的誤差稱為訓練誤差(empirical error)，對測試集數據的誤差稱為泛化誤差(generalization error)，模型對訓練集以外樣本的預測能力就稱為模型的泛化能力。

過擬合(overfitting)是指模型學習能力太強把學習進行的太徹底，捕捉到訓練集本身的特徵，也捕捉到了由於噪聲帶來的假特徵，以至於模型的泛化能力和預測辨識率突然下墜到谷點，圖 4.4 為過擬合現象產生的波形。



圖 4.4 過擬合現象

### 4.3 訓練之混淆矩陣及預測辨識率分析

訓練完成並且沒有發生過擬合時，則可以停止訓練，一般來說訓練迭代次數到 2000 時模型就該停止訓練，預測辨識率也差不多趨近平穩(如圖 4.5 所示)，若持續訓練下去則有可能發生過擬合現象(如圖 4.4)。



圖 4.5 迭代次數到 2000 趨近平穩

訓練完成後會出現兩張混淆矩陣，分別為訓練跟測試(如圖 4.6 及圖 4.7 所示)，我們可以藉由混淆矩陣去瞭解某種雜草誤判為其他雜草有幾張，藉此經由下次訓練前加以更正，以達到更高的辨識率。

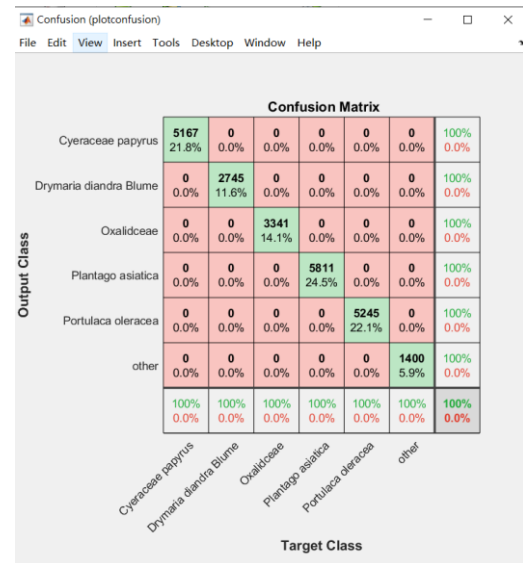


圖 4.6 訓練混淆矩陣

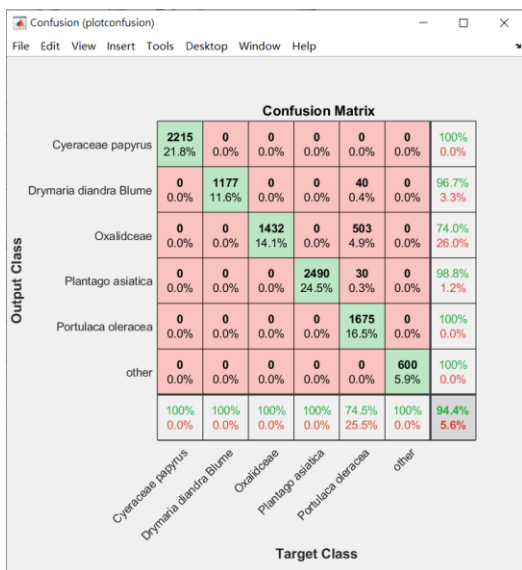


圖 4.7 測試混淆矩陣

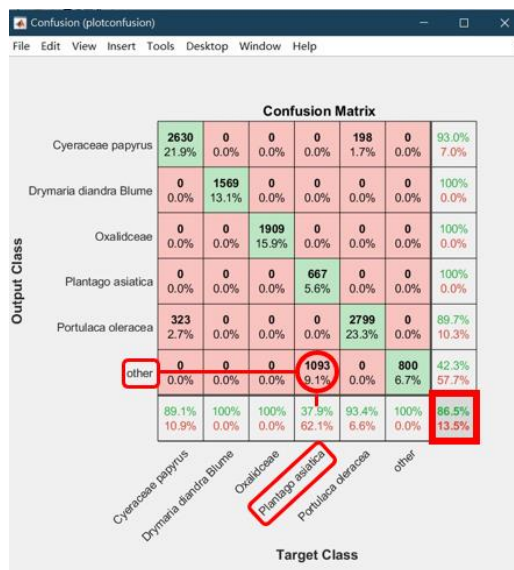


圖 4.8 車前草錯誤判斷

一開始我們使用 VGG-19 神經網路架構進行訓練，訓練完成後在測試混淆矩陣出來時，我們可以看到圖 4.8 車前草(Plantago asiatica)判斷成構樹(other)有 1093 張，正確判斷的照片只有 667 張，單一辨識率只有 37.9%，總辨識率為 86.5%。於是我們增加車前草的 data、換個 ResNet-50 神經網路架構再去訓練一次，結束後我們可以由圖 4.9 得知車前草(Plantago asiatica)完全正確判斷，單一辨識率為 100%，總辨識率來到 94.4%；我們也可以得到因為模型架構的不一樣，所以也會導致其他雜草原本辨識率為 100%的也會差生判斷錯誤。

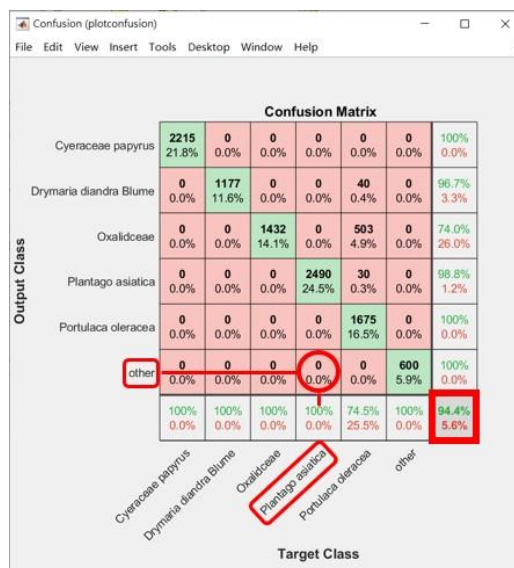


圖 4.9 車前草正確判斷

#### 4.4 辨識結果分析

訓練完成且混淆矩陣沒有太大的錯誤時，就可以輸入照片進行辨識，圖 4.10 為正確判斷的結果、圖 4.11 為誤判的結果。

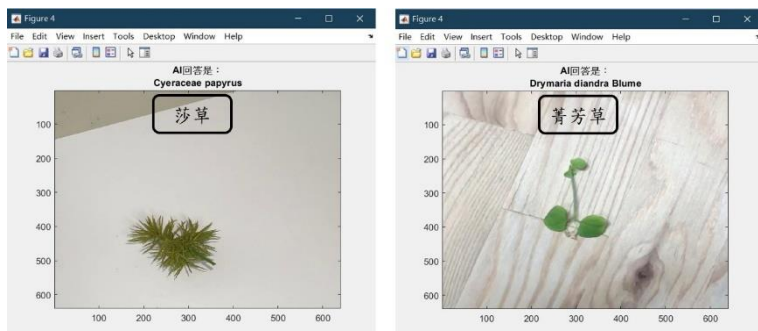


圖 4.10 正確判斷結果

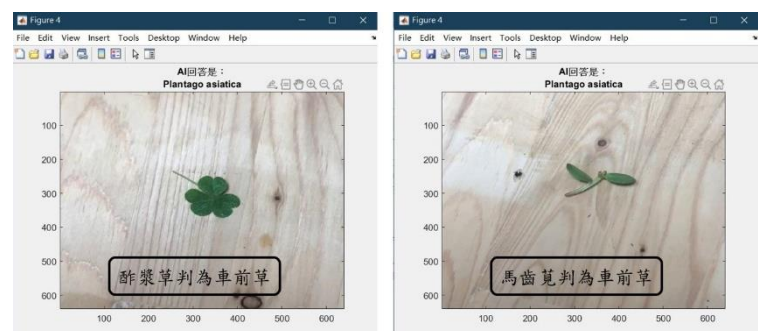


圖 4.11 誤判結果

#### 4.5 土壤溼度感測器結合人機介面成果展示

我們一方面將待驗證之雜草圖片輸入至訓練好的模型，並藉由電腦去判別為何種雜草，一方面透過土壤溼度感測器，將感測器插入該雜草之土壤藉由 arduino UNO 板燒錄程式驅動土壤溼度感測器測得其溼度並且將數值傳送於人機介面中，方便農夫了解土壤的水質多寡。圖 4.12 為土壤濕度感測器之人機介面。

在人機介面中我們設有數值顯示器、土壤溼度情況燈號、濕度波形圖。

- 數值顯示器可以顯示土壤目前濕度數值(0%~100%)
- 土壤溼度情況燈號我們設為三個，分別為貧瘠(0%~30%)、正常(31%~69%)及濕潤(>70%)，這可以馬上瞭解目前土壤的溼度情況
- 濕度波形圖可以瞭解長時間的土壤溼度波形，進而去判斷土壤有沒有異狀

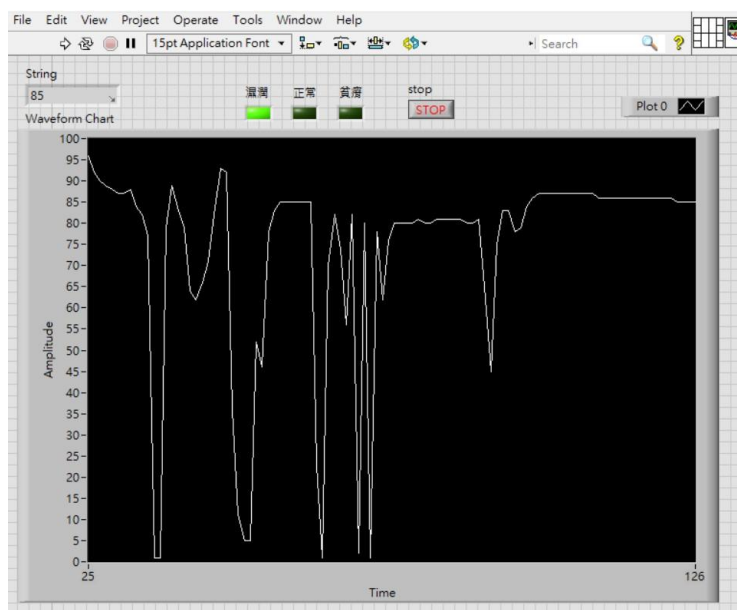


圖 4.12 土壤濕度感測器之人機介面



## 4.6 農園產業結合無人機未來發展預估分析

農園產業結合無人機可以帶來很大的效益，我們可以使用無人機去進行農園視察將無人機所拍攝到的即時影像傳至雲端，類似監視器的概念去查看農園是否有異狀，亦或是可以進行無人機即時影像分析，辨識蔬果成熟度、本專題的雜草辨識等等，這樣不但可以減輕農民的辛苦，也因為現在是農業 4.0 的時代，我們要漸漸將感測器、智能裝置、物聯網、大量資料等等導入農業形成「智慧農園」，以提升農業的效率及帶來更大的產值。

## 五、結論與未來展望

### 5.1 結論

本專題『以無人機進行生態型市民農園六種雜草影像分析及土壤情況分析』藉由 MATLAB 所提供的附加工具箱訓練模組來建構模型，並辨識出雜草種類，透過套用不同的模組、調整迭代次數及學習率，來提高圖片之辨識率，而在製作專題的一開始，我們遇到的問題不外乎就是資料的收集，及圖片解析度的改善，而資料量的多寡將會影響模型辨識出來的效果，因此我們跑遍了許多農園在拍攝雜草，使得資料量達到一定的張數。

接下來利用土壤感測器將土壤中所測得之水分，將數值傳至人機介面上，並利用圖表波形的方式將數值顯示出來。

在製作專題過程中，我們了解到要做到一個完美的模型，其實真的是不太容易，需要一步步的去嘗試每種模型並加以做調整，而近年來物體偵測在 AI 領域蓬勃發展，想必未來會運用在更多領域上面。

### 5.2 未來展望

台灣近幾年來 AI 的興起，從最初的 AlexNet 演變成現今最流行的 YOLO，卷積神經網路變得越來越複雜，但相對的辨識率也有所提升，如再配合無人機即時回傳數據，即時去了解為何種雜草，並加以摘除，即可減輕農民的負擔。

辨識率提升方法有許多種，透過參數調整、神經元的增加及輸入影像的調整，皆可使模型變得更加準確，藉由此次的專題我們也更加了解深度學習的應用及模型的演變及調整，未來將希望能把技術做得更加深入，應用在農業上面，使得農業更佳的自動化，也能讓農民變得更輕鬆。

## 六、參考文獻

- [1] 蔡淳晴,“應用機器學習機制於物件影像辨識之研究-以 TensorFlow 為例”,國防大學資訊管理學系研究所論文,2019
- [2] 劉邦彥,“類自主機器學習與捲積加速器之研究與合成”,國立成功大學電機工程學系研究所論文,2021
- [3] 林季穎,“基於深度學習之人臉特徵辨識與應用”,國立交通大學光電工程學系研究所,2018
- [4] 蔡宗軒,“深度學習應用於種雞辨識之研究”,國立中興大學生物產業機電工程學系研究所論文,2020
- [5] 顏任佑,“提高已知卷積神經網絡精度的有效策略”,靜宜大學資訊工程學系研究所論文,2018
- [6] 王炫尹,“基於深度學習之 PCB 瑕疵偵測”,國立暨南國際大學資訊工程學系研究所論文,2019
- [7] 鄧友豪,“基於遷移式學習之深度學習影像分類參數微調系統”,國立臺灣科技大學電機工程系研究所論文,2018
- [8] A.F.IBRAHIM, S.P.RISTIAWANTO, C.SETIANINGSIH AND B.IRAWAN, “MICRO-EXPRESSION RECOGNITION USING VGG19 CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE AND RANDOM FOREST”, IN *IEEE MICRO-EXPRESSION RECOGNITION USING VGG19 CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE AND RANDOM FOREST*, 2021, pp. 150-156.2